

# Constitutional interests in the face of innovations: how much do we need to know about risk preferences?

Ulrich Witt · Christian Schubert

Published online: 13 June 2008  
© The Author(s) 2008

**Abstract** In constitutional political economy, the citizens' constitutional interests determine the social contract that is binding for the post-constitutional market game. However, following traditional preference subjectivism, it is left open what the constitutional interests are. Using the example of risk attitudes, we argue that this approach is too parsimonious with regard to the behavioral foundations to support a calculus of consent. In face of innovative activities with pecuniary and technological externalities in the post-constitutional phase, the citizens' constitutional interests vary with their risk preferences. To determine what kind of social contract is generally agreeable, specific assumptions about risk preferences are needed.

**Keywords** Constitutional preferences · Social contract · Original position · Innovation · Risk preferences

**JEL codes** D72 · H10 · O33

## 1 Introduction

In order to assess the legitimacy of constitutional rules from an individualistic perspective, Constitutional Political Economy makes frequent use of the device of the social contract. This device structures the comparative normative analysis of alternative constitutional rules, yielding conclusions as to which of them are conducive to mutually beneficial cooperation in the post-constitutional (market) game. Individuals participating at the process of constitutional deliberation are assumed to be situated such that they are induced to favor universal ("fair") over discriminatory rules. The rules stipulated in the social contract are then essentially

---

U. Witt (✉) · C. Schubert  
Max Planck Institute of Economics, Jena, Germany  
e-mail: witt@econ.mpg.de

determined by the “common constitutional interests” of the individuals concerned. Due to the fact, however, that most contributions to Constitutional Political Economy adhere to the methodological premise of “radical subjectivism” (Buchanan and Vanberg 2002), these interests are not materially specified. Rather, in general, a formal interpretation is advanced according to which “‘improvement’ is strictly defined in terms of what the individuals concerned themselves regard as “improvement” (Vanberg 2005, p. 26). Accordingly, the “good” is defined in an endogenous, procedural way, viz., as “that which emerges from agreement among free men, independently of intrinsic evaluation of the outcome itself” (Buchanan 1975, p. 167).

We will argue that this radical subjectivist stance (which corresponds to an analogous methodological position adopted in mainstream welfare economics) considerably weakens the operational power of Constitutional Political Economy. This is particularly true, if it is assumed that, after the constitution has been adopted, agents will play a market game in which innovations are possible. In face of an inherently risky novelty-generating game, individuals will reasonably anticipate that at any time some members of society will be risk-averse (and that they themselves will likely become more risk-averse over time), and their calculus of consent will reflect this. In such a situation, radical subjectivism on constitutional interests and preferences prevents practically relevant and normatively plausible (conjectural) policy conclusions. In order to be able to assess, from the viewpoint of the constitutional stage behind a “veil of ignorance”, the likely impact of future commercial, institutional, or technological novelty on individual income, well-being, and the opportunities for mutually beneficial exchange, additional behavioral hypotheses are needed.

We will support our claim by drawing upon hypotheses about the individuals’ risk attitudes and the systematic changes in them over an individual’s life-span. This somewhat more “objective” (behavioral) approach would, we will argue, allow theorists interested in evaluating innovations on an individualistic basis to use the device of the (hypothetical) social contract for their purposes. Hence, our approach implies an enlargement of the informational base of the contractarian approach to welfare which may be seen as being analogous to the one advocated by scholars such as Sen (1979) in the realm of orthodox, quasi-utilitarian welfare economics. In the same sense as Sen’s contributions, we suggest to follow the creed that “[t]he account of well-being must be tailored to the human frame: it must connect with what makes human lives go well, not the lives of ideal persons.” (Qizilbash 1997, p. 263)

The paper is organized as follows. Section 2 surveys the way the social contract model is usually employed in Constitutional Political Economy. Distinguishing between so called “pecuniary” (market-mediated) external effects of innovations on the economic conditions of the citizens and “technological” (not market-mediated) externalities, we discuss in Sect. 3 the importance of making specific assumptions on individual risk attitudes first in regard to the pecuniary externalities. It turns out that the risk of negative pecuniary externalities—often deemed normatively irrelevant in Constitutional Political Economy (see, e.g., Vanberg 2006)—does affect the terms of the social contract. Section 4 elaborates on the analogous

calculus of consent for the case of technological externalities. It is shown that unanimous consent is facilitated by including specific regulations on innovativeness in the social contract. Section 5 offers our conclusions.

## 2 Different versions of the contractarian argument in constitutional political economy

As the contemporary heir to classical political economy, modern Constitutional Political Economy is essentially based on two conceptual pillars: First, it can be defined as applying “the methods...of modern economics to the study of the basic rules under which social orders may operate” (Brennan and Hamlin 1998, p. 401). Thus, it starts from the fundamental methodological distinction between the level of the “rules of the market game” and the level of single exchange acts within a given system of rules. It is focused on examining both the behavioral effects and the normative legitimacy of alternative rules by arguing that, due to their complexity and open-ended character, market outcomes can only be influenced “indirectly” by shaping the institutional arrangements that bring them about (Vanberg and Buchanan 1991). This focus on rules we will dub the *instrumental* position in the following.

Second, constitutional political economy is characterized by a distinctly *normative* position which is based on the principle of “normative individualism”. According to this principle, only the real individuals’ preferences or interests are accepted as genuine sources of value (Buchanan 1990).<sup>1</sup> It is presumed that a generally agreeable, supra-individual (“objective”) source of normativity is absent so that it is only the voluntary bilateral exchange between consenting individuals which can be claimed to be able to generate legitimacy. The objects of exchange remain unspecified, as do the individuals’ preferences. Related to this subjectivist creed, Constitutional Political Economy elaborates upon the Wicksellian idea of extending the concept of economic efficiency from the market sphere to the study of collective choice processes by applying the procedural logic of exchange acts, centering around the subjectively perceived *gains from trade* that turn up as gains from joint commitment in the realm of constitutional choice (Buchanan 1977; Buchanan 1991a).<sup>2</sup> As a consequence, the normative position held by Constitutional Political Economy is grounded on the device of the *social contract* that generalizes this idea to the level of collective (multilateral) exchange processes. The social contract is used as a tool to assess the legitimacy of specific constitutional rules, by allowing to test whether rational

<sup>1</sup> Vanberg (2005, pp. 24–26) defines it, somewhat broader, as “the presumption that the welfare of the individuals concerned represents the relevant standard against which market and state are to be judged” and as prescribing that “the individuals involved must themselves be respected as the ultimate judges on what qualifies as ‘good’ or desirable in their social transactions and rule arrangements”.

<sup>2</sup> “[I]ndividuals choose to impose constraints on their own behavior...as a part of an exchange in which the restrictions on their own actions are sacrificed in return for the benefits that are anticipated from the reciprocally extended restrictions on the action of others...” (Buchanan 1990, p. 4).

individuals would be willing to accept these rules, when placed in a “constitutional choice” situation.

The two conceptual premises have to be carefully distinguished. In principle, the instrumental focus on rules would also be compatible with an outcome-oriented utility maximization paradigm (with “outcomes” defined negatively or in pattern-based form). As regards the normative focus, the consensus-oriented gains from trade-paradigm could also be applied—practical difficulties notwithstanding—to alternative interventions at the level of single market transactions and outcomes. Let us have a closer look at the two pillars.

As to the first one, the focus on more or less abstract rules governing social interaction is usually justified by the fact that the post-constitutional market game can best be characterized as an open-ended evolutionary process, being constituted and driven by creative, inventive choices on the part of (at least some of) the individual market participants (Buchanan and Vanberg 2002). Consequently, specific future states and outcomes cannot be predicted, let alone deliberately be brought about. Since it is assumed that no value scalar exists independently of the choices made, the efficiency of outcomes can only be inferred in terms of the characteristics of the process from which they result, rather than in terms of the “welfare attributes” of the outcomes per se (Vanberg 2005). In other words, “the basis for any normative... judgment must be shifted from patterns of results to characteristic features of the institutional-structural parameters themselves” (Buchanan and Vanberg 2002, p. 124). Only rules and procedures can be “improved” in a meaningful sense. This hypothesis is derived from the epistemological assumption that, as a result of, say, cumulative “trial-and-error-learning” (Vanberg 2005, p. 37; see also Vanberg 1994), technological knowledge about the working properties of rules is more easily available than the much more specific knowledge about particular welfare effects of outcome-oriented interventions. Hence, it is assumed that even if detailed statements about the future course of events are impossible, the general working properties of alternative rules can be examined, as can the overall pattern of costs and benefits that they may generate over time. The resulting insights can then be used to inform choices among alternative constitutional regimes.

As to the second pillar, viz., the individualistic position on normativity, the use of the social contract plays a key role. It is somewhat more intricate. The historical predecessors of modern constitutional political economy—Hobbes in particular—had used this concept in order to demonstrate the idea that from the viewpoint of a situation of general social anarchy, it would be in the interest of everyone involved to enter into a general agreement to refrain from using violence against others and to accept a centralized state agency—the notorious *Leviathan*—in order to violently enforce the respective set of regulations. Put differently, everyone would reasonably assent to the multilateral exchange of behavioral constraints that is implied by such a mutual disarmament. By tapping general agreement as a source of legitimacy in its own right, Hobbes was able to move, in an “individualistic” spirit, beyond traditional objective approaches to normativity.

While retaining their predecessors’ meta-ethical stance on norms and values, modern contractarians pursue more modest aims. Rather than deriving the

legitimacy of the state in general, they use the social contract model to test the normative appeal of specific constitutional rules and to develop plausible statements as to their normative quality. This is reflected in the Pareto criterion or, equivalently, the criterion of *unanimity* which is often (wrongly, as it will turn out) held to be the only normative premise which the contractarian argument requires. Unanimity can be seen as the only criterion that is able to properly reflect the principle of normative individualism in the realm of collective choice. It is epitomized by the subjectively perceived mutual gains that people can reap when they exchange behavioral constraints by jointly committing themselves to rules (Buchanan 1991b, 81ff). These gains determine the “common constitutional interests” that constitute the only normative measuring rod which most authors in constitutional political economy accept.

The unanimity test is carried out by way of a thought experiment about how individual agents would decide within some artificially specified constitutional situation, the “original position”. Hence, modern contractarians use the social contract device in a *hypothetical* sense, if often only implicitly. This is justified by the fact that, due to prohibitive transaction costs (“costs of collective decision-making” in Buchanan and Tullock’s 1962 terminology), real consent is obviously unattainable in a complex society. If it were actually attainable, contractarians would not need to be consulted anyway. Using hypothetical agreement only implies, however, that the conclusions obtained cannot claim to be more than mere conjectural recommendations. For as already Hume (1748/1992) emphasized in his “On the Original Contract”, fictitious contracts can only effectively bind fictitious individuals.<sup>3</sup> Consequently, any demonstration that some given rule would be agreeable to all individuals participating at a hypothetical vote on some social contract does not logically imply that the same rule would be agreed upon by real individuals under real circumstances. It does not carry with it genuine legitimizing force. Thus, such a demonstration can at best serve as one possible (ideally psychologically effective) contribution to the public deliberation that those real individuals concerned by the rule are engaged in when reasoning about and discussing alternative constitutional arrangements.

Such a contribution implies hypothetical imperatives, discerning which rules should turn out to be generally acceptable for the individuals, given some plausible hypothesis on those individuals’ constitutional interests. Nonetheless, it is the real consent of these individuals only in terms of which the legitimacy of any proposed rule can be judged (a point emphasized by Vanberg 2005 and Buchanan 1991a, b among others). Hence, a careful distinction is required between the social contract’s role as a *normative tool* (possibly one among many others) that may be based on more or less individualistic premises, and its role as a model for actual (ideally perfectly individualistic) procedures that eventually confer legitimacy on constitutional rules.

<sup>3</sup> See also Buchanan (1977), Kliemt (1986), Müller (2002) and Binmore (1994, p. 39), who argues that “(t)here is nothing about the circumstances under which people are hypothesized as bargaining in the original position that can justify the assumption that they are bound, either morally or in practice, to abide by the terms of a hypothetical deal reached in the original position.” A similar argument has however already been made by Hobbes (see Hardin 1998, p. 647).

Given these clarifications about the status of contractarian conclusions, how exactly are these conclusions generated? Methodologically, the contractarian argument works like a deductive-nomological prediction (see Müller 2000, pp. 37–46, referring to Hempel and Oppenheim 1948). Given some universal law and a set of antecedent conditions, a singular statement (the *projectandum*) can be logically deduced. In the context of a contractarian argument, the universal law is represented by some behavioral hypothesis, usually a strict variant of the rational choice model. The set of antecedent conditions is represented by some assumptions about the individuals' utility functions, their informational endowments or their relative bargaining power. Individuals can, e.g., be assumed to aim at maximizing their personal utility, be mutually disinterested, be uncertain about their future position in society, but on the other hand to possess perfect technological knowledge about the working properties of alternative constitutional rules. From the behavioral hypothesis and the situational assumptions a (more or less general) statement follows logically as to what kind of rule those fictitious individuals would choose.

It is evident that any modification of the situational assumptions, say, about the set of admissible individual preferences, may alter the contractarian conclusions in a potentially fundamental way by shaping the set of possible welfare-increasing states that are accessible from the original position. That is why designing these assumptions involves (often only implicitly) genuine value judgments. There is consequently a wide variety of contractarian approaches that differ in terms of precisely those assumptions that concern both what individuals *want* and what they *know* in the original position. Following Sugden (1993), it is convenient to distinguish between “strong” and “weak” versions of the social contract model. The main difference concerns the respective theorist's ambition to purge the whole approach from any undesired positive (in particular psychological) and evaluative presuppositions deemed superfluous.

As to the strong variant, it is best exemplified by Buchanan's (1975) attempt to legitimize the modern liberal “rule of law” state with its protective and productive branches as emerging from a general consent among individuals living under conditions of a rule-less (hence, Pareto-inferior) “natural anarchy”.<sup>4</sup> The exact conditions, including in particular the individuals' initial endowments, that prevail in this original position are deliberately left open, for any attempt to define them would invoke specific value judgments, and Buchanan wants to keep his approach “value-free” except for the unanimity criterion itself. Thus, according to this version of contractarianism, anything is deemed legitimate that individuals turn out to agree upon. Hence, this approach can be described as being based upon a “strongly individualistic” notion of justice (Müller 2000, pp. 49–50). As the conditions under which they agree are left unspecified, even agreements obtained under circumstances implying an extremely unequal bargaining power (such as a slave's acceptance of her master's rule) carry legitimizing force; consequently, Buchanan cannot exclude the possibility that a “fair” social contract would include slavery.

<sup>4</sup> Another case of a strong contractarian approach is Gauthier's (1986) rational bargaining-based model of morality.

Due to the low level of specification, the resulting social contract approach has an extremely low degree of operational content. It cannot move beyond the observation that what is unanimously chosen by the individuals is Pareto-superior to the original position and Pareto-optimal. No practical policy conclusions can be derived, as they would involve a choice between alternative Pareto-optima. Needless to say that *a fortiori*, this approach cannot “predict” the emergence of a complex liberal rule-of-law state, as Buchanan attempted to. Notice that the strictly Paretian approach to Welfare Economics suffers from exactly the same lack of operational content: Since the Pareto criterion can only define a “strict partial ordering” among the set of possible social states, it cannot discriminate between different points located on the Pareto frontier (Sen 1970, pp. 21–30). Again, the lack of operational content is due to the fact that only one, extremely abstract, value judgment is admitted to the analysis.

As regards the weak versions of the social contract model, they attempt to restrict the set of possible conclusions (“predictions”) by constraining the set of admissible conditions within the original position. Only those constitutional choices are deemed to be “fair” that satisfy certain procedural fairness requirements. This principle is usually realized by introducing a “veil of uncertainty” (Buchanan and Tullock 1962) or “veil of ignorance” (Rawls 1971) that serves as a metaphor for artificially restricting the information available to the individuals in the original position. It is based on the presupposition that, when deliberating about alternative constitutional rules from behind a sufficiently “thick” veil, even rational and strictly self-interested individuals are forced to occupy a “moral” viewpoint, where “moral” is to be understood as implying universalisability. They are induced to judge rules from an impartial position, i.e., to figure out how they would evaluate a given rule if they would be in any of the social positions that are taken to be available in the post-constitutional stage. This then increases the chances that even self-interested individuals may agree upon a set of rules that truly are in their common “constitutional interest”.

Within the set of weak social contract models, there is a second distinction to be taken into account between those approaches that interpret the veil (and the original position more generally) in an *empirical* way and those construing it essentially as a *normative* device. As regards the empirical variant, it is based on the assumption that in constitutional decision-making situations, individuals typically deliberate on rules that are relatively general, abstract and durable. Hence, their precise effects on particular welfare positions are hard to predict. This is a fact that prevents individuals from judging those rules from their own, narrowly conceived self-interest. It leads them to favor exactly those general rules that constitute the “just” constitutional market order that benefits everyone involved (Brennan and Buchanan 1985). Relatedly, Vanberg (2005, p. 31, italics added) sees the use of the veil construct as serving to clarify the question “what kinds of procedures for choosing rules are more likely than potential alternative procedures to result in the adoption of rules that are ... mutually beneficial,” stipulating that “conjectures of this kind are about the *factual* working properties of procedures...for choosing rules rather than about subjective evaluations of the persons concerned.” Based on these insights, hypotheses can be advanced as to how the “thickness” of the veil should be



deliberately manipulated in order to increase “the prospect of achieving agreement” (Vanberg and Buchanan 1989, p. 54). While a thicker veil makes agreement on values (“interests”) more probable, consensus on positive constitutional “theories” will be facilitated by making the veil sufficiently “thin” (ibid; see, however, Müller 1998 for a qualification of this assumption). Insights such as these would then serve as the basis for policy advice concerning the design of collective decision-making procedures.

For good reasons,<sup>5</sup> most contractarian approaches use the social contract in a normative sense, i.e., as a device that explicitly embodies a given set of value judgments. In particular, the veil of uncertainty or ignorance is employed as representing and effectively introducing some version of a moral impartiality requirement, by forcing even selfish individuals to adopt a neutral viewpoint with respect to the constitutional rules they are bargaining or deliberating upon. As Kliemt (2004) emphasizes, the original position needs to be modeled in such a way that it is “embedded” in a sufficiently rich ethical background in order to allow for meaningful and determinate conclusions. This approach implies, however, first, that the ambition to base constitutional advice on *purely* individualistic grounds is abandoned. Due to its extremely thin evaluative basis, accounts based on pure Normative Individualism prove unable to narrow the range of admissible conclusions sufficiently to avoid indeterminacy. External value judgments may also be necessary to avoid strong counterintuitive results, such as Buchanan’s (1975) slavery contracts: The proposal by Vanberg (1986) to supplement Buchanan’s strictly individualistic unanimity criterion by introducing a (necessarily non-individualistic) cost-based criterion of “voluntariness” that excludes any involuntarily concluded contracts from the evaluative space of his contractarian theory, exemplifies this normative approach. The second implication is that the ambition to develop a normative device that is universally applicable has to be abandoned, for the ethical basis that determines the original position’s antecedent conditions will necessarily reflect the moral common sense of a particular cultural background (Pettit 1974). All this does however not pose a problem as long as the social contract model is not used as a legitimizing device in the sense explained above, but only as a normative tool.

These methodological presuppositions favor a weak (veil-centered) and normative social contract approach which, however, is still short in operational relevance. A way to make progress in this respect is to enrich the contractarian model with material hypotheses about human behavior as the backdrop of the individuals’ constitutional interests. If the individuals anticipate to be playing a post-constitutional innovative market game, a straight forward example of such hypotheses are the human attitudes towards risk. Although the discussion of a behavioral foundation for constitutional political economy must be confined here to this example, we claim that

<sup>5</sup> Evidently, apart from its questionable empirical assumptions (constitutional deliberation often involves quite specific rules) this approach runs into problems of circularity: it presupposes exactly what it purports to explain. When in the constitutional choice set only “general” rules are available, it is trivially clear that “general” rules will be agreed upon in the end. It may be argued, therefore, that the interpretation of the social contract (exclusively) in empirical terms is a nonstarter (Müller 2000, pp. 80–81).



such a foundation is a desideratum more generally. For if the antecedent conditions of the contractarian argument are left unspecified, the model can only generate an extremely thin partial ordering of social states and rules that does not allow to make any substantive statements about the content of the social contract.

It is true that, on the basis of the criteria of unanimity and voluntariness alone, the radical subjectivist stance on constitutional interests still arrives at a conclusion: from behind the “veil”, individuals will unanimously agree to a set of constitutional rules that maximize their chances to engage in mutually beneficial exchange on the basis of whatever preferences they may have (Vanberg 2006). Such a conclusion would recommend all institutional arrangements that enlarge the scope of decentralized “systems competition”. But even this conclusion already presupposes a specific preference for certain “basic goods” (such as liberty and individual responsibility) over others which cannot be taken for granted on empirical grounds (see Buchanan 2005). We do not deny the plausibility of the above conclusion for certain areas of constitutional rule-making; we just question its status as the only object of unanimous consent from behind a “veil of ignorance”. In order to reach more concrete (hypothetical) statements as to the normative quality of constitutional rules, the traditional radical subjectivism on individual interests is not helpful.

### **3 Constitutional preferences and risk attitudes: the case of pecuniary externalities of innovations**

In this section we will discuss the need for more specific hypotheses about the preferences of the citizens when it comes to the calculus of consent about the terms of the social contract. More specifically, we focus here and in the next section in an exemplary fashion on the role that risk attitudes play for determining the interest component of the constitutional preferences (Vanberg and Buchanan 1989) of the participants in the social contract. The significance of risk attitudes results, we will argue, from the fact that human inventiveness and innovative action are characteristic features of the post-constitutional market game. The “discovery procedure” in which the market process results (Hayek 1978) does not necessarily produce outcomes that are to everyone’s advantage. However, what forms the market game can take depends on the rules stipulated in the social contract concerning innovativeness, in short called the “innovation regime” here.

As explained elsewhere in detail, the regime may allow more or less innovativeness and/or discriminate between different forms of innovations, i.e. invoke more or less, and different forms of, regulations on innovations (Witt 1996). How the social contract is specified in this respect cannot be decided without making some assumptions about the citizens’ risk attitudes. Hence, the need for more specific hypotheses about these attitudes. To lay out the argument, it is useful to first discuss in generic terms the consequences of market innovations and the potential risks they imply. Since the seminal work of Schumpeter (1934, 1942), innovativeness and innovations are seen as something beneficial and, hence, desirable in economics, an assessment that is now widely shared in public. The background of the Schumpeterian interpretation is the historical experience of a positive correlation between

innovativeness on the one hand and economic growth on the other. Indeed, for the past two centuries, the correlation between innovations and a rising per capita income seems evident (Rosenberg and Birdzell 1986; Mokyr 1990).

However, innovations by definition mean newly implemented problem solutions which have not been tried before. For this reason, who profits from the beneficial effects of an innovation hinges not only on whose problems the innovation solves, but also on whether that problem solving triggers new problems elsewhere. There are obvious risks implied here. First, a rising per capita income usually does not (fully) reflect the social costs that innovations can incur on society (or on some of its members) in the form of negative technological externalities. Innovations can have unforeseen harmful effects on health and/or the environment. Second, it is only on average that innovativeness results in a growth of per capita income. Innovations regularly also produce losers among some participants in the market game—a fact Schumpeter (1942) already alluded to with the euphemistic formulation “creative destruction”. These effects in the form of negative pecuniary externalities usually have to be born unevenly.

In view of risks like these implied by a post-constitutional market game with unregulated innovativeness some interesting questions arise. A first (normative) question is whether technological and pecuniary externalities should be considered equally relevant for the calculus of consent at the constitutional stage. A second (methodological) question is whether the calculus of consent regarding these risks should be based on material hypotheses about human risk attitudes. Following the considerations in the previous section, both questions will be answered in the affirmative here. The reason regarding the first question is that, in the eyes of the citizens who deliberate on the stipulations of the social contract behind a veil of ignorance, it does not matter whether a possible destitution to be faced later is technologically caused or caused by market forces. In both cases the citizens have to strike a balance between benefits and risks of innovations which they consider acceptable. The reason for the affirmative answer to the second question is that speculations about (hypothetical) calculi of consent that abstract from fundamentals of human nature seem rather artificial.

Concerning the citizens' information about what will happen in the post-constitutional market game, we assume that they do not know whether they will have a chance to become an innovator themselves and how innovation(s), be they their own or those of other market participants, will affect their well-being. With respect to the risk attitudes it seems reasonable to assume that the citizens not only know their own attitudes, but are also aware of the fact that risk preferences or, more generally, attitudes towards perceiving and taking risks, show a significant, inter-personal variance between individuals (a point also emphasized by Vanberg 2006). The empirical variance reflects genetically determined differences, cultural influences, and the state of individual experiences. Moreover, independent of the inter-personal variance at each point in time, risk preferences systematically change over time with increasing age. On average there is a greater inclination to take risks among juveniles and young, particularly male, adults. Beyond early adulthood, risk aversion grows with increasing age (see Rubin and Paul 1979 and the references given there).

For this reason it is almost certain that, both at the constitutional stage and in the post-constitutional market game, there will be participants who are already highly risk-averse or become so over their further life cycle. Given that these facts about risk attitudes are part of human nature, it seems straightforward that, in the calculus of consent at the constitutional stage, they are recognized. Two questions then arise: whose risk attitudes will determine the stipulations of the social contract regarding innovativeness, and what kind of stipulations will this be? Keeping to the unanimous consent criterion, each citizen has a veto right. Since citizens with more risk friendly preferences tend to tolerate more innovation risks, they always strike a more favorable balance between benefits and risks of innovations than less risk friendly agents. A (strategically motivated) veto by the former would therefore not be credible—it would hurt them more in their perception than it would hurt the risk-averse. Given that the balance between benefits and risks appears least favorable to the most risk-averse citizen, only this agent can exert a credible threat so that the stipulations of the social contract regarding innovativeness will hinge on what is agreeable to her or him.

Our approach differs in several respects from related contractarian models proposed by Vickrey (1945, 1960) and Harsanyi (1953, 1975) on the one hand and Rawls (1971) on the other hand. First of all, besides our emphasis on innovativeness, our model of the constitutional decision-making calculus does not include a representative agent, but rather a multitude of heterogenous agents. As regards the seminal accounts by Vickrey and Harsanyi (who introduced the expected utility calculus to model the choices of a representative “impartial observer” in the original position), we diverge from their assumptions about the “preference space”. Both Vickrey and Harsanyi assume that the representative agent expects to occupy any of the conceivable post-constitutional positions with equal probability. This key postulate (based on the contentious Laplace rule) then leads to the prediction that the agent will rank outcomes in a utilitarian way. The conclusion critically hinges, however, on the fact that both authors refrain from imposing any structure on the individuals’ preference space. While Vickrey takes agents to assess future positions according to their given constitutional preferences, Harsanyi assumes them to abstract from their own tastes and (using empathetic “extended preferences”) to put themselves into the shoes of possible future selves. In contrast, we assume that in both the constitutional and the post-constitutional setting, the preference space is not completely open. Rather, some agents have and/or will have risk-averse preferences for certain.

As regards Rawls’ (1971) model, this does make specific assumptions concerning the agents’ risk attitudes—an approach we share. But, unlike Rawls who takes the agents to be perfectly ignorant about their *future* risk preferences, we assume some knowledge about human nature, i.e., about the general patterns of how these preferences develop in the course of one’s life time. Furthermore, we claim that risk attitudes show a significant inter-personal variance, while Rawls assumes that behind the veil the agents are identically (and to an extreme extent) risk-averse. This assumption notwithstanding, Rawls considers the agents to be unable to form subjective estimates for the likelihood of occupying any conceivable position in the post-constitutional game, including the worst position. In contrast, we assume that

agents are able to do this (see below). Accordingly, we do not share Rawls' view that the agents uniformly stick to the maximin criterion when judging alternative sets of rules and corresponding patterns of outcomes.

Since the stipulations of the innovation regime need to be differentiated with respect to the kind of potential negative externalities and the ways of coping with them, we focus in this section on market-mediated pecuniary externalities first. (For the discussion of technological externalities neglected here see the next section.) From his contractarian perspective Viktor Vanberg writes on the negative pecuniary externalities: "The very productivity of the 'game of the market' depends on its openness for the discovery of 'superior' products and production, and in agreeing to play the market game people accept the risk that their investments may be devalued by such discoveries. The productive advantage of the market game and this risk are two sides of the same coin. One cannot have the one without the other, *and fairness in playing the market game requires all players to accept this risk*. As the theory of rent-seeking shows, players may, of course, lobby governments for special privileges that allow them, at the expense of other players, to enjoy the benefits of the market's productivity while avoiding its accompanying risks" (Vanberg 2006, p. 207, italics added). Does such an assessment do justice to the underlying problem of risk preferences differing between the citizens?

Imagine an innovation that raises productivity and/or increases the quality of a commodity or service produced by one of the competitors in a market. If the demand side accepts the innovation (i.e. if the innovative offer is able to make customers better off), other competitors in the market regularly face problems for their business. They lose customers and orders, often with the consequence that labor has to be laid off. Expected returns from specific investments made earlier when the innovation was not yet known have to be given up for lost. This can mean losses of private wealth for capital owners or foregone expected income from human capital investments, if not even unemployment and loss of current income for members of the labor force. Under freedom of contract, the gains from an innovation succeeding in the market are always larger for the economy as a whole than what the competitors lose by the pecuniary externalities (see Witt 1987)—this being the very reason for the historical increases of average per capita income. Nonetheless, not knowing who will be winners and who will be losers of future innovations, the fact that such losses have to be born by some agents is highly relevant for risk-averse citizens who later have to accept the related stipulations in a social contract.

Let the citizens be indexed  $i = 1, \dots, n$ . For expository convenience assume that, at the constitutional stage, i.e. behind a veil of ignorance, the calculus of consent can be represented as follows. The expected outcome of a "laissez faire" innovation regime (indicated by suffix lf) is given by a simple lottery  $L(y_{lf}, p)$  in which an income  $y$  is earned with probability  $p$ , both incomes and probabilities being subjective estimates of the citizens (index suppressed subsequently). More specifically, the participants in the post-constitutional market game suffering from negative pecuniary externalities of the successful innovations of others are assumed to expect a low income  $y_l$ . In contrast, the income of those who do not expect to suffer the externalities can be assumed to have a significantly higher value  $y_h$ . Let the expectation of being a player in the post-constitutional market game who suffers

from pecuniary externalities be given by the probability  $p$  and the expectation of the opposite event by  $1-p$ . Hence, the income a member of society expects is

$$E(y_{lf}) = py_l + (1 - p)y_h = y'_{lf}, \quad (1)$$

where  $y'_{lf}$  is the same amount of money obtained with certainty. Risk aversion then means that the utility  $u(E(y_{lf}))$  derived from the expected income is strictly smaller than the utility  $u(y'_{lf})$  derived from the sure income  $y'_{lf}$ . Put the other way round, for citizen  $j$  with the most risk averse preferences the inequality

$$CE(L(y_{lf}, p))_j < y'_{lf} \quad (2)$$

holds, where  $CE(L(y_{lf}, p))_j = c_j$  denotes  $j$ 's certainty equivalent of the lottery characterizing the “laissez faire” innovation regime, i.e. a sure income  $c_j$ .

Now imagine the only alternative innovation regime would be the opposite of unconstrained innovativeness, namely a regime requiring unanimous approval of any innovation by all parties affected by its possible negative pecuniary externalities. Abstracting from transaction costs and/or strategic maneuvering that may in practice hamper such a regime, there would be no innovations causing negative pecuniary externalities either because they would be vetoed or because those profiting from the innovation would have to pay compensations to the losers to obtain their approval. Since nobody would be made worse off, but some may be better off, by innovations, this regime can be dubbed a “Pareto” innovation regime (indicated by suffix  $p$ ).<sup>6</sup> Compared to the “laissez faire” regime, a “Pareto” innovation regime considerably lowers the benefits accruing from innovative activities—and hence the incentives to undertake them. In fact, in the presence of transaction costs and/or strategic maneuvering, the “Pareto” regime is likely to result in a complete breakdown of innovative activities. All existing competitors would then be protected and the dynamic competitiveness of the markets would be grossly reduced, as would be technical progress and economic growth.

However, unlike under the “laissez faire” regime where gains and losses may be accumulating in a vastly uneven way among the citizens, there will be nobody suffering pecuniary externalities where others enjoy the high incomes that innovations make possible. Behind the veil of ignorance, all citizens can expect from a “Pareto” regime a low, but sure income  $y_p$ . Of course, compared to the expected income under the “laissez faire” regime the condition  $y_p < E(y_{lf})$  holds. For a rational decision making on the stipulations of a social contract at the constitutional stage it then follows:

**Proposition 1** *At the constitutional stage, the most risk-averse citizen  $j$  consents to a social contract that stipulates a “laissez faire” innovation regime, if and only if*

$$CE(L(y_{lf}, p))_j \geq y_p. \quad (3)$$

<sup>6</sup> The “Pareto” regime implicitly presupposes that only those innovations are carried out that result in compensation payments not exceeding the income constraint(s) of the innovator(s), see Kerber (1993) for a discussion.

To prove this, assume contrary to (3) that  $CE(L(y_{lf}, p))_j < y_p$ . Consenting to a “laissez faire” regime under that condition would mean that citizen  $j$  reveals a preference for a sure income  $c_j < y_p$ . Since no other motives have been assumed this would violate the rationality assumption. Hence, a rational risk-averse agent  $j$  consents to a “laissez faire” regime rather than a “Pareto” regime if and only if condition (3) is satisfied.

The “laissez faire” regime and “Pareto” regime may be extreme cases, but they are not unrealistic. A *grosso modo* positive (laissez faire like and, hence, risk friendly) attitude towards innovations as it is characteristic for the public opinion nowadays seems to be a rather recent phenomenon. For most of human history, public attitudes towards innovations were much less supportive of, if not hostile towards, innovativeness, at least in the sphere of production and trade. Indeed, there are historical examples of regulations that strongly resemble a “Pareto” regime. The guild systems in Europe and Asia, for instance, denied newcomers free access to the local markets (North 1981, Chap. 10, Jones 1988, Chap. 7). Higher authorities usually backed the enforcement of market closure so that potential newcomers had little, if any, room for mustering political power to fight their exclusion (cf. Ekelund and Tollison 1981, Chaps. 3 and 4). The institutional set-up of at least the European guild systems, which had a guild assembly as the collective decision making body, meant that collective consent was also necessary for innovations by guild members. Often unanimous approval was required before innovations could be introduced and practiced. Innovative cross-fertilization between different professional and/or regionally distinct guilds substantially bridled.<sup>7</sup>

“Laissez faire” and “Pareto” regimes are, of course, not the only possibilities. Indeed, with somewhat more complex innovation regimes it is even possible to soften condition (3) in proposition 1 without imposing constraints on innovativeness in the post-constitutional market game. One such possibility is, for instance, to combine a laissez faire regime with a tax financed social security system that guarantees to those suffering the pecuniary externalities an income never falling short of the subsistence level  $y_{min}$ . For such a “laissez faire with social security” regime (indicated by suffix  $s$ ) the following can be established:

**Proposition 2** *Under certain conditions, the most risk-averse citizen  $j$  who is unwilling to consent to a “laissez faire” innovation regime, is willing to consent to a social contract that stipulates a “laissez faire with social security” innovation regime.*

To demonstrate the conditions under which proposition 2 holds let us explore the “laissez faire with social security” regime more closely. Assume that a transfer payment  $t = \max \{y_{min} - y_l, 0\}$  is guaranteed by the government so that a citizen’s income will at least be  $y_{min}$  with certainty. Let  $t$  be financed by a tax  $T$  invoked on those not suffering the pecuniary externalities, i.e. on high income earners, where the tax revenue is always set equal to the size of the transfer payments (balanced

<sup>7</sup> It accords with its hostility towards, and a low pace of, innovativeness that the medieval European economy basically stagnated over the period 1500–1820, see Maddison (2001, Tables B.19 and B.22).

budget condition). Hence,<sup>8</sup>  $T = [p/(1 - p)]t$ . The expected income then changes into:

$$E(y_s) = E(y_{lf}) + pt - (1 - p) T. \quad (4)$$

Note that under the balanced budget condition, the expected incomes  $E(y_s) = E(y_{lf})$ . Since the expected income in excess of the sure subsistence level is given by

$$E(y_s) - y_{\min} = p \max \{y_1 - y_{\min}, 0\} + (1 - p)(y_h - y_{\min}) - [p/(1 - p)] \max \{y_{\min} - y_1, 0\} \quad (5)$$

the right hand side of Eq. 5 represents a lottery  $L(y_{ex}, p)$  of its own. Accordingly, the certainty equivalent of the lottery  $L(y_s, p)$  for citizen  $j$  can be expressed as a sum

$$CE(L(y_s, p))_j = y_{\min} + CE(L(y_{ex}, p))_j > CE(L(y_{lf}, p))_j. \quad (6)$$

Despite the equality in the expected values of the income in Eq. 4, the inequality sign regarding the certainty equivalents in (6) holds, because of the sure income component  $y_{\min}$  guaranteed by the government. Consequently, it is possible that the condition

$$CE(L(y_s, p))_j \geq y_p > CE(L(y_{lf}, p))_j$$

holds. Indeed, this is the necessary condition for proposition 2 to hold. In that case the most risk-averse member of society consents to a “laissez faire with social security” regime while, by condition (3) in proposition 1, that member will not consent to a “laissez faire” innovation regime.

As it appears, the question of what stipulations of the social contract regarding innovativeness can find unanimous consent depends on what is assumed about the specific risk preferences of the citizens. Different assumptions lead to strongly differing stipulations. From the fact that in many democracies a majority of voters show a rather consistent tendency towards endorsing social security transfer payments one may infer that these votes reveal a risk aversion. Many of these payments are indeed contingent on the occurrence of risks that are related to negative pecuniary externalities of large scale innovativeness.<sup>9</sup> Taken at face value, the majority thus seems to express a political preference for something like a “laissez faire with social security” innovation regime.<sup>10</sup> Yet, the preferences revealed by voting in the *post*-constitutional phase are one thing and the hypothetical reconstruction of the decision that would have been made behind the

<sup>8</sup> For analytical convenience it is assumed that  $y_h > T$ .

<sup>9</sup> Their historical background has been the extreme inequality in the distribution of the innovation benefits during the industrial revolution, exposing large fractions of the population to pauperism, famines, and an extremely low life expectancy, see, e.g., the report in Engels (1968).

<sup>10</sup> If so, this would explain the majority vote despite the fact that the benefactors of the social security transfer payments form a comparatively small group endowed with little voting power. The alternative explanation in public choice theory, which interprets the social security transfer payments as a result of inefficiency causing rent-seeking activities (see Mueller 1993), has to refer to hypotheses that do not *prima facie* seem more plausible. It is claimed, e.g., that voters are negligent of the small burden placed on them by each single interest group privilege and therefore accept it. Or it is conjectured that the majority vote results from logrolling between many rent-seeking interest groups in the political process.



veil of ignorance another one. The very purpose of the constitutional choice calculus is to escape from the many biases expressed in factual voting turnouts due to partisan interests and strategic voting. Taking recourse to voting turnouts is therefore no option for deriving hypotheses about risk preferences.

We have argued that the diversity of risk attitudes and their systematic change over a life time are part of human nature. On the basis of this behavioral hypothesis it can be conjectured that, at all times, there are citizens who are already strongly risk-averse or anticipate to become so later. If this conjecture is true, it is likely that, under the unanimity rule, only rather protective innovation regimes for the post-constitutional market game will find consent. This may either be a regime that compensates for negative pecuniary externalities resulting from “creative destruction” by some kind of social security payments—a result that is not easily accommodated with the usual identification of the concept of “social security” with rent-seeking activities of interest groups. Or this may be an innovation regime like the “Pareto” regime that rules out such externalities, together with a large share of potential innovativeness.

#### **4 Constitutional preferences and risk attitudes regarding technological externalities of innovations**

Let us turn now to the constitutional preferences of the citizens regarding negative technological externalities of innovations and the corresponding question of what stipulations of the social contract are ‘consentable’. To ease the analysis we will now abstract from the existence of pecuniary externalities. While in the case of negative pecuniary externalities where, under freedom of contract, the sum of private gains from an innovation being carried out always exceed the sum of private losses, in the case of negative technological externalities of an innovation this is different. The social costs can by far exceed the sum of the private gains of the agents profiting from an innovation. Moreover, these excessive costs often only turn out with a considerable time delay and may only then alert the affected agents or the society as a whole to the social costs that have to be borne.

Examples of such developments easily come to mind. In the 1950s asbestos was introduced into the markets as innovative material that was emphatically welcome where incombustible or fireproof products were needed. In the late 1970s it was discovered that dust particles from asbestos have a serious carcinogenic impact in humans. Another case is fluorocarbon. It was introduced and widely adopted as multi-purpose innovation serving as a lubricant, refrigerant, fire-extinguishing agent, insulator, and aerosol propellant. But then evidence was gradually emerging that it played a causal role in the Earth’s stratospheric ozone layer depletion. Or take the example of nuclear power that was propagated, publicly welcomed, and heavily subsidized as the major innovation in energy technology in the 1960s. Its hazards were drastically experienced only later in power plant accidents in Three-Mile-Island and Chernobyl.

In terms of Coase’s (1960) classical formulation of the problem of social costs, the inevitable risk of yet unknown negative technological externalities of

innovations means that there are inherent, non-negligible transaction costs. These make it difficult, if not impossible, to rely on *ex ante* negotiations between an innovator who would cause the externality and those who would be affected as a way of efficiently internalizing the social costs. The social costs arising from innovations at later points in time, e.g., from damages to health or the environment, are therefore an object necessitating publicly enforced regulations. In the spirit of constitutional political economy, their agreeability can be assessed by comparison with the stipulations of a social contract that would find unanimous consent by the citizens at the constitutional stage.

By the same logic as in the case of pecuniary externalities, different stipulations in the social contract lead to different innovation regimes. These, in turn, are associated with differences in the risk and/or the extent of the damages. Since money income is no adequate measure for environmental and health damages, the negative externalities will be taken into account here by their effects on subjective well-being or utility. The best way in which regulations can cope with the harmful effects is to prevent them. Accordingly, innovation regimes can now be distinguished by the means they use to prevent innovations from developing negative technological externalities and the extent to which they succeed in doing so. One means would be a ban on (or injunction of) innovations—an “innovations ruled out” regime that suppresses technological externalities by suppressing innovations in the first place. (In a legal perspective, it is equivalent to granting every member of society a non-attenuated property right in an externality-free environment.) This regime can serve here as a benchmark for discussing the implications of different assumptions about risk preferences.

Other regimes use liability as a means (i.e. attenuate the mentioned property right). Such rules cannot make harmful effects and the corresponding losses in well-being undone. But they can create incentives for the innovator to recognize potential social costs as early and as seriously as possible and to internalize them in her or his decision making. Hence, liability rules can influence both the likelihood and the size with which social costs arise. Such regimes do not rule out innovations in toto, but differ with respect to how far they hold the innovator liable for any negative technological externalities. As will be explained in a moment in more detail, they range from a “no liability” regime at the one end via a “limited liability” regime to a “strict liability” regime at the other end of the spectrum.

Which of these regimes can be expected to (hypothetically) find unanimous consent at the constitutional stage, and under what conditions, has now to be clarified. Again there will be citizens who are, or anticipate to become, highly risk averse, and it is the consent of the most risk averse member to one of the innovation regimes that needs to be assured for the unanimously approved social contract. The calculus of consent can be modeled along similar lines as the risk preference model of the Eqs. 1–6. Yet, the different innovation regimes have to be assessed now by the impact they have on the level of well-being that can be reached as a result of innovations on the one hand and the loss of well-being due to their negative technological externalities on the other hand.

Let the level of well-being attainable by innovations that have no negative technological externalities be  $w_1$ . The level of well-being resulting if innovations do

cause damages can be denoted by  $w_I - w_D$ . Again these variables and the probabilities for a damage to occur or not to occur,  $q$  and  $(1 - q)$  respectively, are taken to be the subjective estimates of member  $i$  (index suppressed subsequently). The expected outcome of the innovation regimes can then be represented by a simple lottery  $L(w, q)$ . In the benchmark case of an “innovations ruled out” regime the lottery degenerates to a level of well-being  $w_{\min}$  that, innovations and their potential hazards being absent, is taken to result with certainty.

In the light of the examples given above it is realistic to assume that the characteristics and the technological externalities of innovations are not instantaneously recognized in full. They are rather revealed only gradually. Those agents closer involved with the innovation are likely to notice problems earlier than others. The levels of well-being that a citizen can expect to attain when innovations are allowed is then likely to vary with the liability rules. Let the expected value under a “no liability” innovation regime (indicated by suffix  $nl$ ) be given by<sup>11</sup>

$$E(w_{nl}) = q(w_I - w_D) + (1 - q)w_I. \quad (7)$$

Since in the absence of any liability the innovator is given no incentive to care about negative technological externalities of her or his innovation (except to the extent of the share she or he has to bear), it is not very likely that she or he will take much effort to (promptly) respond to emerging evidence of innovation-induced damages.<sup>12</sup> For this reason, the loss of well-being,  $w_D$ , that the citizens have to bear can become very large. Indeed, it can become so large that the expected contribution of an innovation to individual well-being turns negative.

This would be different if innovators are made liable in a strict and unlimited way for the social costs their innovations turn out to cause, i.e. under a “strict liability” innovation regime (indicated by suffix  $sl$ ). A strong incentive would be given to proceed cautiously, to use emerging information as carefully as possible, and to react in a prompt and determined manner once externalities are discovered. Accordingly, a lower damage to well-being  $w_D' < w_D$  would result. However, all care taken cannot prevent negative technological externalities to occur that, by the nature of novelty, cannot be anticipated at the time of carrying out the innovation. Even when stopped instantaneously, damages cannot be made undone. An innovator who is held liable, strictly and without limits, for any ill effects that may later be discovered evidently faces an uncontrollable risk of a potentially devastating loss. For any kind of risk taking attitude the threat of excessive damage claims will negatively impact the willingness to carry through an innovation. Yet with a significantly reduced innovativeness, the level of individual well-being which the citizens can reach as a result of innovations will also be reduced. Let the new level be  $w_I' < w_I$ . This reduction can overcompensate the damage reduction. If the

<sup>11</sup> For a level of well-being  $w_{nl}'$  attained with certainty for which  $w_{nl}' = E(w_{nl})$  holds, risk aversion then again implies that  $u(E(w_{nl})) < u(w_{nl}')$ .

<sup>12</sup> In fact, it can even be argued that there is an incentive for the innovator to withhold a reaction in order to induce other citizens to pay her or him for discontinuing the innovation (see Schlicht 1996).

expected value for the level of well-being under the “strict liability” innovation regime is denoted by:

$$E(w_{sl}) = q(w'_I - w'_D) + (1 - q)w'_I, \quad (8)$$

it is thus possible that  $E(w_{sl}) < E(w_{nl})$ .<sup>13</sup>

How can such an overcompensation be avoided without removing the incentive to react appropriately when negative externalities turn up? One way would be to delimit the innovator’s liability to only those damages that he or she could have prevented by immediately taking appropriate measures, when the first hints to a damaging effect are discovered. Such a limited liability implies a principal-agent problem. Unlike the innovator, the public in the role of the principal is incompletely informed about both the occurrence of externalities and timely remedies. However, ex post the efforts taken by the innovator in the role of the agent can be observed and insufficient performance can be sanctioned. A credible threat of ex post sanctions creates incentives to take appropriate effort. The institutional solution to the principal-agent problem is, in legal terms, to apply a liability for negligence rule to the mastering of discovered externalities.

An institutional solution such as this means that the risk of unforeseen damages is not born by the innovator alone, but in part also by society as a whole. Assuming that the conditions to which her or his liability extends are indeed under the control of the innovator, such an innovation regime does not reduce the innovative efforts. Unlike under the “strict liability” regime, the level of well-being  $w_I$  will therefore not be reduced. As under the “strict liability” regime, however, the damage to well-being will be reduced to  $w'_D$ . Hence, the expected value for the level of well-being under a “limited liability” innovation regime in the specified sense (indicated by suffix ll) is given by

$$E(w_{ll}) = q(w_I - w'_D) + (1 - q)w_I. \quad (9)$$

In view of the Eqs. 7–9 it is obvious that, in terms of the levels of attainable well-being, the lottery  $L(w_{ll})$  corresponding to the “limited liability” innovation regime dominates the lotteries  $L(w_{sl})$  and  $L(w_{nl})$  corresponding to the other two regimes. Accordingly, for all citizens independent of their risk preferences

$$CE(L(w_{ll}, q)) > CE(L(w_{sl}, q)) \text{ and } CE(L(w_{ll}, q)) > CE(L(w_{nl}, q)). \quad (10)$$

Turning now to the calculus of consent of the most risk-averse citizen  $j$  regarding the negative technological externalities occurring only after innovations have been carried through the following can be shown to hold:

**Proposition 3** *The most risk-averse citizen  $j$  consents to a social contract stipulating an innovation regime other than “innovations ruled out”, if and only if it is a “limited liability” regime for which*

$$CE(L(w_{ll}, q))_j \geq w_{\min} \quad (11)$$

<sup>13</sup> Inserting Eqs. 7 and 8, this condition holds if  $(w_I - w'_I)/(w_D - w'_D) > q$ .

The first claim in proposition 3 follows from the order relation (10). To prove the second claim assume that, contrary to condition (11),  $CE(L(w_{II}, q))_j = v_j < w_{\min}$ . Consenting to the “limited liability” regime under that condition would mean that member  $j$  reveals a preference for a sure level of well-being  $v_j$  smaller than the sure level  $w_{\min}$  that can be attained under the “innovations ruled out” regime. Since no other motives have been assumed this would violate the rationality assumption. Hence, the rational most risk-averse agent  $j$  consents to a “limited liability” innovation regime if and only if condition (11) is satisfied.

To enforce a “limited liability” innovation regime a society must be able to create and support the necessary legal institutions. What one observes, however, are also publicly financed facilities for doing systematic research on potential hazards of existing and currently developing technologies. Since the corresponding knowledge that is being created has the features of a public good, its provision through contract research privately financed by an innovator would likely fall short of what is desirable. Moreover, research of this kind may be used by a publicly financed health agency or an environmental agency to implement state-of-the-art protective measures *ex ante*. By public authority these may be enforced as a standard by making innovations contingent on the agency’s approval.<sup>14</sup> The probability of doing harm to health and/or the environment implied by innovations can thus be reduced. This does not necessarily reduce the economic benefits, nor does it, of course, mean that the yet unknown potential amount of damage to well-being will be lower, should a non-anticipated negative externality nonetheless occur. For expository purposes, the well-being variables  $w_I$  and  $w_D'$  can therefore be taken to be unaffected. However, the research and regulation activities come at a cost to the public that needs to be financed by raising a tax from each citizen which causes a corresponding reduction in well-being  $w_T$ .

Let us dub an innovation regime with these extensions a “limited liability with regulatory agency” regime (indicated by suffix  $lr$ ). If the subjectively expected hazard reduction is represented by a probability  $\varepsilon$ ,  $0 < \varepsilon < q$ , so that  $q - \varepsilon = r$ , the expected value of well-being associated with such a regime is

$$E(w_{lr}) = r(w_I - w_D') + (1 - r)w_I - w_T = E(w_{II}) + \varepsilon w_D' - w_T. \quad (12)$$

On the basis of the assumptions underlying Eq. 12 it follows:

**Proposition 4** *If condition (11) is not satisfied and the most risk-averse citizen  $j$  is therefore unwilling to consent to a “limited liability” innovation regime,  $j$  will consent to a “limited liability with regulatory agency” innovation regime, if the certainty equivalent of the hazard reduction exceeds  $w_T$  sufficiently.*

To prove proposition 4 consider the calculus of consent of the rational, risk-averse citizen  $j$ , who does not consent to the “limited liability” regime, because  $CE(L(w_{II}, q))_j < w_{\min}$  (proposition 3). This citizen will not consent to a “limited liability with regulatory agency” innovation regime either, if the certainty equivalents of the two lotteries representing the innovation regimes satisfy the condition  $CE(L(w_{lr},$

<sup>14</sup> An example is the admission procedure for new drugs practiced by the U.S. Food and Drug Administration. The approval by public authorities (which are usually risk-averse) can be seen as a control against an overly risk-taking behavior of innovators regarding their liability.

$r))_j \leq CE(L(w_{ll}, q))_j$ . In view of Eq. 12 this is exactly the case, if the certainty equivalent of  $\varepsilon w_D'$  is smaller than, or equal to, the sure amount  $w_T$ . Hence, for consent as stated in proposition 4 it is necessary, but not sufficient, that the difference between the certainty equivalent of the hazard reduction and the sure amount  $w_T$  is positive. The sufficient condition is that the difference becomes so large that

$$CE(L(w_{lr}, r))_j \geq w_{\min} > CE(L(w_{ll}, q))_j$$

holds. This is the more likely, the less risk-averse  $j$  is, the greater  $j$  assesses the hazard reduction that is possible, and the lower the agency costs  $w_T$ .

As in the previous section, the two propositions derived here demonstrate that a legitimization of some particular innovation regime by means of a hypothetically reconstructed consent to corresponding stipulations of the social contract is only possible on the basis of specific hypotheses about the citizens' risk preferences. Again we have drawn here on the risk-related behavioral hypothesis derived from human nature that predicts the presence of citizens with strongly risk-averse preferences. If they do not even favor an "innovations ruled out" regime, unanimous consent is likely to result in something like a regime involving publicly financed agencies that regulate technological innovations—an institution sometimes considered incompatible with the idea of "a minimal state" in Constitutional Political Economy.

In the developed world, something like a "no liability" regime was prevailing during the early phases of industrialization. It was gradually turned into a "limited liability" regime and, with rising per capita real income, a slow conversion to a "limited liability with regulatory agency" regime developed. In more recent times, in some of the most prosperous countries even political preferences for partial "innovations ruled out" regimes become apparent, e.g. with respect to innovations in gene technology or cloning techniques. From the fact that a rising share of the citizens favor regulations, or even a partial ban, of innovations in their voting behavior it may again be inferred that there is considerable risk aversion. Yet, as already emphasized at the end of the previous section, these votes reflect *post*-constitutional preferences. They are not cast behind a veil of ignorance and can therefore not be accepted as unambiguous indicators of the true constitutional interests of the citizens.

## 5 Conclusions

In this paper we have attempted to show that its adherence to a strong form of preference subjectivism prevents Constitutional Political Economy from developing more definite statements about the normative quality of alternative constitutional innovation regimes. By a constitutional innovation regime we understand the stipulations of a hypothetical social contract regarding the conditions under which, in the post-constitutional market game, innovations are legitimate. It turned out to be implausible to presume that individuals, when situated behind a "veil of ignorance" with respect to the positions they will hold in the post-constitutional market game, will unanimously opt for a *laissez faire* regime, even when only

pecuniary externalities are considered in isolation. Given their knowledge about their own risk attitudes and about the general human disposition to become more risk-averse in the course of one's life, they can rather be argued to opt for a regime that is supplemented by a social security net. In the case of negative technological externalities risk-averse preferences have been shown to favor a limited liability rule supplemented by publicly financed regulatory agencies. Both findings were reconstructed here from the rational calculus of the most risk averse citizen participating in constitutional decision-making.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

## References

- Binmore, K. G. (1994). *Game theory and the social contract. vol I: Playing fair*. Cambridge, MA: MIT Press.
- Brennan, G., & Buchanan, J. M. (1985). *The reason of rules*. Cambridge, MA: Cambridge University Press.
- Brennan, G., & Hamlin, A. (1998). Constitutional economics. In P. Newman (Ed.), *The new palgrave dictionary of economics and the law*, vol. I (pp. 401–410). London: Macmillan.
- Buchanan, J. M. (1975). *The limits of liberty*. Chicago: University of Chicago Press.
- Buchanan, J. M. (1977). A contractarian paradigm for applying economic theory. In J. M. Buchanan (Ed.), *Freedom and constitutional contract* (pp. 235–242). College Station: Texas A&M University Press.
- Buchanan, J. M. (1990). The domain of constitutional economics. *Constitutional Political Economy*, 1, 1–18.
- Buchanan, J. M. (1991a). The Foundations of Normative Individualism. In Ders, *The economics and ethics of constitutional order* (pp. 221–231). Ann Arbor: University of Michigan Press.
- Buchanan, J. M. (1991b). *The economics and ethics of constitutional order*. Ann Arbor: The University of Michigan Press.
- Buchanan, J. M. (2005). Afraid to be free: Dependency as desideratum. *Public Choice*, 124, 19–31.
- Buchanan, J. M., & Tullock, G. (1962). *The calculus of consent*. Ann Arbor: University of Michigan Press.
- Buchanan, J. M., & Vanberg, V. J. (2002). Constitutional implications of radical subjectivism. *Review of Austrian Economics*, 15, 121–129.
- Coase, R. H. (1960). The problem of social cost. *Journal of Law and Economics*, 3, 1–44.
- Ekelund, R. B., Jr., & Tollison, R. D. (1981). *Mercantilism as a rent-seeking society: Economic regulation in historical perspective*. College Station: Texas A&M University Press.
- Engels, F. (1968). *The condition of the working class in England*. Stanford: Stanford University Press.
- Gauthier, D. (1986). *Morals by agreement*. Oxford: Oxford University Press.
- Hardin, R. (1998). Modern contractarianism. In P. Newman (Ed.) *The new palgrave dictionary of economics and the law*, vol. II (pp. 645–651). New York: Stockton Press.
- Harsanyi, J. C. (1953). Cardinal utility in welfare economics and in the theory of risk taking. *Journal of Political Economy*, 61, 434–435.
- Harsanyi, J. C. (1975). Can the maximin principle serve as a basis for morality? A critique of John Rawls' theory. *American Political Science Review*, 69, 594–606.
- Hayek, F. A. (1978). *New studies in philosophy, politics, economics and the history of ideas*. London: Routledge.
- Hempel, C. G., & Oppenheim, P. (1948). Studies in the logic of explanation. *Philosophy of Science*, 15, 135–178.
- Hume, D. (1748/1992). Of the original contract. In T. M. Green & T. M. Grose (Eds.), *David Hume-the philosophical work* (pp. 443–460). Aalen: Scientia.



- Jones, E. L. (1988). *Growth recurring—economic change in world history*. Oxford: Clarendon.
- Kerber, W. (1993). Rights, innovations and evolution: the distributional effects of different rights to innovate. *Review of Political Economy*, 5, 427–452.
- Kliemt, H. (1986). Individualism, libertarianism and non-cognitivism. *Analyse & Kritik*, 8, 211–228.
- Kliemt, H. (2004). Contractarianism as liberal conservatism: Buchanan's unfinished philosophical agenda. *Constitutional Political Economy*, 15, 171–185.
- Maddison, A. (2001). *The world economy: A millennial perspective*. Paris: OECD.
- Mokyr, J. (1990). *The lever of riches: Technological creativity and economic progress*. Oxford: Oxford University Press.
- Mueller, D. C. (1993). *The public choice approach to politics*. Edward Elgar: Aldershot.
- Müller, C. (1998). The veil of uncertainty unveiled. *Constitutional Political Economy*, 9, 5–17.
- Müller, C. (2000). *Das vertragstheoretische Argument in der Ökonomik*. Berlin: Duncker & Humblot.
- Müller, C. (2002). The methodology of contractarianism in economics. *Public Choice*, 113, 465–483.
- North, D. C. (1981). *Structure and change in economic history*. New York: Norton.
- Pettit, P. (1974). A theory of justice. *Theory and Decision*, 4, 311–324.
- Qizilbash, M. (1997). Needs, incommensurability and well-being. *Review of Political Economy*, 9, 261–276.
- Rawls, J. (1971). *A theory of justice*. Cambridge, MA: Belknap Press.
- Rosenberg, N., & Birdzell, L. E. (1986). *How the west grew rich: The economic transformation of the industrial world*. New York: Basic Books.
- Rubin, P. H., & Paul, C. W. (1979). An evolutionary model of taste for risk. *Economic Inquiry*, 17, 585–596.
- Schlicht, E. (1996). Exploiting the coase mechanism: The extortion problem. *Kyklos*, 49, 319–330.
- Schumpeter, J. A. (1934). *The theory of economic development*. Cambridge, MA: Harvard University Press.
- Schumpeter, J. A. (1942). *Capitalism, socialism, and democracy*. New York: Harper.
- Sen, A. K. (1970). *Collective choice and social welfare*. Amsterdam: North-Holland.
- Sen, A. K. (1979). Personal utilities and public judgments: Or what's wrong with welfare economics? *Economic Journal*, 89, 537–558.
- Sugden, R. (1993). Rationality and impartiality: Is the contractarian enterprise possible? In D. Gauthier & R. Sugden (Eds.), *Rationality, justice and the social contract* (pp. 157–175). Ann Arbor: University of Michigan Press.
- Vanberg, V. J. (1986). Individual choice and institutional constraints. *Analyse & Kritik*, 8, 113–149.
- Vanberg, V. J. (1994). Cultural evolution, collective learning and constitutional design. In D. Reisman (Ed.), *Economic thought and political theory* (pp. 171–204). Boston: Kluwer.
- Vanberg, V. J. (2005). Market and state: The perspective of constitutional political economy. *Journal of Institutional Economics*, 1, 23–49.
- Vanberg, V. J. (2006). Human intentionality and design in cultural evolution. In C. Schubert & G. von Wangenheim (Eds.), *Evolution and design of institutions* (pp. 197–212). London: Routledge.
- Vanberg, V. J., & Buchanan, J. M. (1989). Interests and theories in constitutional choice. *Journal of Theoretical Politics*, 1, 49–62.
- Vanberg, V. J., & Buchanan, J. M. (1991). Constitutional Choice, Rational Ignorance and the Limits of Reason. *Jahrbuch für Neue Politische Ökonomie*, 10, 61–78.
- Vickrey, W. S. (1945). Measuring marginal utility by reaction to risk. *Econometrica*, 13, 319–333.
- Vickrey, W. S. (1960). Utility strategy, and social decision rules. *Quarterly Journal of Economics*, 74, 507–535.
- Witt, U. (1987). How transaction rights are shaped to channel innovativeness. *Journal of Institutional and Theoretical Economics*, 143, 180–195.
- Witt, U. (1996). Innovations, externalities and the problem of economic progress. *Public Choice*, 89, 113–130.